

CLAIMS

What is claimed is:

1. A method of mining association rules from datasets while maintaining privacy of individual transactions within said datasets through randomization, said method comprising:
 - randomly dropping true items from each transaction;
 - randomly inserting false items into each transaction; and
 - estimating the nonrandomized support of an association rule in the original dataset given its support in the randomized dataset.
2. The method in claim 1, wherein said randomization comprises per transaction randomizing, such that randomizing operators are applied to each transaction independently.
3. The method in claim 1, wherein said randomization is item-invariant such that a reordering of said transactions does not affect outcome probabilities.
4. The method in claim 1, wherein said dropping of said true items and said inserting of said false items are carried out to an extent such that the chance of finding a false itemset in a randomized transaction relative to the chance of finding a true itemset in said randomized transaction is above a predetermined threshold.
5. The method in claim 4, wherein said predetermined threshold provides that the chance of finding a false itemset in said randomized transaction is approximately equal to the chance of finding a true itemset in said randomized transaction.
6. The method in claim 1, wherein said dropping of said true items and said inserting of said false items are performed independently on said transactions prior to the transactions being

collected in the database.

7. A method of mining association rules from databases while maintaining privacy of individual transactions within said databases through randomization, said method comprising:
randomly dropping true items from each transaction;
randomly inserting false items into each transaction; and
mining said database for association rules after said dropping and inserting processes by estimating the nonrandomized support of an association rule in the original dataset given its support in the randomized dataset.

8. The method in claim 7, wherein said randomization comprises per transaction randomizing, such that randomizing operators are applied to each transaction independently.

9. The method in claim 7, wherein said randomization is item-invariant such that a reordering of said transactions does not affect outcome probabilities.

10. The method in claim 7, wherein said dropping of said true items and said inserting of said false items are carried out to an extent such that the chance of finding a false itemset in a randomized transaction relative to the chance of finding a true itemset in said randomized transaction is above a predetermined threshold.

11. The method in claim 10, wherein said predetermined threshold provides that the chance of finding a false itemset in said randomized transaction is approximately equal to the chance of finding a true itemset in said randomized transaction.

12. The method in claim 7, wherein said dropping and said inserting are performed independently on said transactions prior to the transactions being collected in the database.

13. A method of mining association rules from datasets while maintaining privacy of

individual transactions within said datasets through randomization, said method comprising:

creating randomized transactions from an original dataset by:

randomly dropping true items from each transaction in said original dataset, and

randomly inserting false items into each said transaction;

creating a randomized dataset by collecting said randomized transactions; and

mining said database for association rules after said dropping and inserting processes by estimating nonrandomized support of an association rule in the original dataset based on the support for said association rule in said randomized dataset.

14. The method in claim 13, wherein said process of creating randomized transactions comprises per transaction randomizing, such that randomizing operators are applied to each transaction independently.

15. The method in claim 13, wherein said process of creating randomized transactions is item-invariant such that a reordering of said transactions does not affect outcome probabilities.

16. The method in claim 13, wherein said dropping of said true items and said inserting of said false items are carried out to an extent such that the chance of finding a false itemset in a randomized transaction relative to the chance of finding a true itemset in said randomized transaction is above a predetermined threshold.

17. The method in claim 16, wherein said predetermined threshold provides that the chance of finding a false itemset in said randomized transaction is approximately equal to the chance of finding a true itemset in said randomized transaction.

18. The method in claim 13, wherein said process of creating randomized transactions is performed independently on said transactions prior to the transactions being collected in said randomized database.

19. A program storage device readable by computer, tangibly embodying a program of instructions executable by the computer to perform a method of mining association rules from databases while maintaining privacy of individual transactions within said databases through randomization, said method comprising:

randomly dropping true items from each transaction;
randomly inserting false items into each transaction; and
mining said database for association rules after said dropping and inserting processes by estimating the nonrandomized support of an association rule in the original dataset given its support in the randomized dataset.

20. The program storage device in claim 19, wherein said randomization comprises per transaction randomizing, such that randomizing operators are applied to each transaction independently.

21. The program storage device in claim 19, wherein said randomization is item-invariant such that a reordering of said transactions does not affect outcome probabilities.

22. The program storage device in claim 19, wherein said dropping of said true items and said inserting of said false items are carried out to an extent such that the chance of finding a false itemset in a randomized transaction relative to the chance of finding a true itemset in said randomized transaction is above a predetermined threshold.

23. The program storage device in claim 22, wherein said predetermined threshold provides that the chance of finding a false itemset in said randomized transaction is approximately equal to the chance of finding a true itemset in said randomized transaction.

24. The program storage device in claim 19, wherein said dropping and said inserting are performed independently on said transactions prior to the transactions being collected in the database.